International Journal of Information and Computer Science
IJICS Volume 1, Issue 2, May 2012 PP. 39-45
ISSN(online) 2161-5381
ISSN(print) 2161-6450
www.iji-cs.org

Acclimatization Calculation Research Based on Multiple Stereovision

Yadan Pei [1] Qiuqi Ruan [2]
Institute of Information Science ,Beijing Jiaotong University,Beijing,China
Email: 10120395@bjtu.edu.cn

(Abstract) As environment factors such as illumination, posture greatly affect the identification results on face recognition, it's a key point for machines to adapt themselves to complex and changing environment. In this paper, a new idea of concept acclimatization is presented, and we take the method of Multiple Stereovision and adaptive selection to solve the problem caused by posture change under the acclimation calculation framework. Also a decision level data fusion method, the combination of General Regression Neural Network and Dempster-Shafer Evidence Theory is presented to help improve accuracy of recognition.

Keywords: Acclimatization; Multiple Stereovision; Data Fusion; General Regression Neural Network; Dempster-Shafer Evidence Theory

1. INTRODUCTION

Face recognition is a research emphasis in image processing, pattern recognition, compared with other biometrics recognition technology, face recognition is more convenient, direct and easily accepted, it is widely used in authentication, video surveillance, criminal investigation and other fields. At present, most algorithms in face recognition are all for standard front face images. Even we ignore factors such as illumination, occlusion, the face images we get may have wide range of changes in the rotation angle, it is difficult to make a correct face recognition.So it is an important direction for us to have a accurate feature points location and recognition under various posture.

Research shows that most methods for posture change are based on monocular vision system, such as multi-view Eigen-Face [1], efficient 3D reconstruction [2].Although these methods can solve the problem in some degree, they need large amount of calculation and complex process. At the same time, changes of external environment is various, there are still many restrictions to solve these problems under monocular model. So we present the multi-vision method, get multiple view face images of one target with the use of multiple cameras. Also we can make use of redundant and complementary information from multi-sensors to improve the recognition rate.

The remainder of the paper is organized as follows. In section 2,acclimation is defined, and the acclimation calculation framework under multi-vision is presented. Section 3 gives a detail introduction of the implementation method of acclimation calculation. Section 4 presents a decision level fusion algorithm which combines General Regression Neural Network with Dempster-Shafer Evidence Theory. Finally section 5 concludes the paper.

2. ACCLIMATION CALCULATION UNDER MULTI-VISION

2.1. Multiple Stereovision System and Acclimation Calculation

Multiple Stereovision is to place multiple cameras to capture multi-view images of the same target, as shown in **Figure 1**, one of the cameras synchronizes with others. We can get more image information about target in Multi-Vision system than in monocular vision, which helps to reduce treatment difficulties, improves the accuracy and robustness of the detection system. Compared with binocular vision, increase in the number of cameras leads to an increase of geometric constraints, which reduces the difficulties of matching.

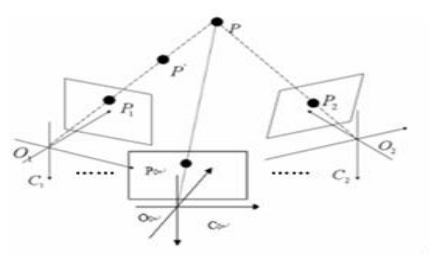


Figure 1. Multiple Stereovision System

In this paper, we define acclimation as the process that machines adapt themselves to external environment by means of detecting key environment parameters and adjusting system parameters automatically. The framework of acclimation calculation based on multiple stereovision mainly

contains three stages, input, pose estimation and acclimation calculation, as in **Figure 2**. In acclimation calculation, three steps are organized, key factor detection, calculation and decision.

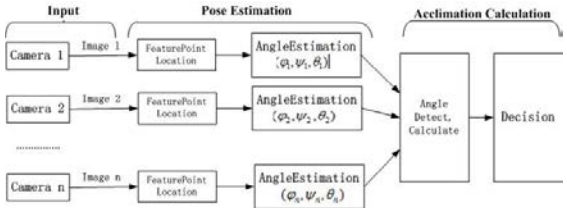
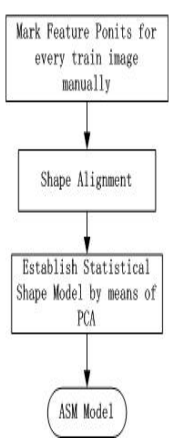


Figure 2. Acclimation Calculation Framework

2.2. Pose Estimation

There are many methods for pose estimation, Choi [3] take 6 feature points to estimate face rotation angle based on EM algorithm, it requires for long-time running. Lei Wang [4]take advantages of multiple cameras to locate 8 feature points and estimate rotation angles, this method needs camera calibration and matching, it's a large computing.Guoyuan Liang [5]use 3D face model and affine correspondences to estimate face pose, it highly depends on the accuracy of 3D model data and it is complicated. For our method presented in the paper, we should first locate 5 feature points, the outer corner of left eye P_{Le} , the outer corner of right eye P_{Re} ,the nose tip P_{Np} , the outer corner of left mouth P_{Lm} , the outer corner of right mouth P_{Rm} , then estimate face pose (φ,ψ,θ) on the basis of projection theory.

1)Face feature location



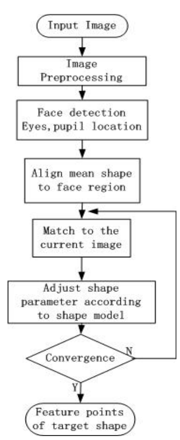


Figure 3. Build ASM model **Figure 4.** Feature points location

We apply ASM(Active Shape Model)algorithm to locate face features. ASM(Active Shape Model)presented in [6] is a statistical model of the shape of objects which iteratively deform to fit to an example of the object in a new image, the shapes are constrained by the PDM(Point Distribution Model)to vary only in ways seen in a training set of labeled examples. The shape of an object is represented by a set of points. The ASM algorithm aims to match the model to a new image. It works by alternating the following steps. First, look in the image around each point for a better position for that point. Second, update the model parameters to best match to these new found positions. **Figure 3** and **Figure 4** show the process of ASM algorithm. **Figure 5** gives a location image of 68 feature points on the basis of ASM algorithm.

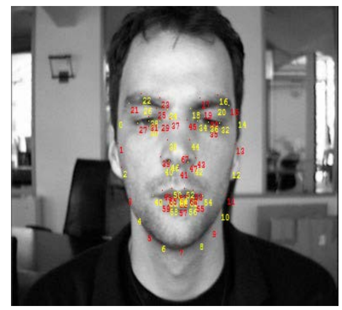


Figure 5. Location of 68 feature points

2)Rotation angle estimation

- Expression of face posture

Methods used to represent face pose are various, for example, use azimuth and dip angle, rotation angle in three-dimensional coordinate system and so on. In this paper, rotation angle are adopted. In rotation angle method, face is treated as 3D rigid body in three-dimensional coordinate system, images with various pose is regarded as 3D rigid body's projection in two-dimensional plane. Face is viewed as symmetric. Define the middle point of the connecting line of left and right outer eye corner O as origin of coordinates for both two-dimensional plane Oxy and three-dimensional coordinate system $OXYZ$,see **Figure 6**.

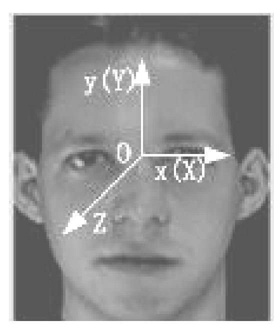


Figure 6. Coordinate system in face image

Without considering displacement of face image, the relation between a point $p(x,y,z)$ in three-dimensional coordinate system and the corresponding one in two-dimensional plane $P(X,Y)$ is computed as **Eq.1**,in which s is scale coefficients, and U is direct projective matrix **Eq.2**.

$$\begin{pmatrix} X \\ Y \end{pmatrix} = sU\Phi \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{1}$$

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad (2)$$

Φ is a transformation matrix. For front face image, Φ is unit matrix, while various face postures have different transformation matrix. According to rotation angle algorithm, various posture are equivalent to the front face's rotation about X , Y and Z axes, which can be presented by a set of different rotation angles (φ, ψ, θ) ,so Φ can also be presented as

$$\Phi = R_1(\varphi) R_2(\psi) R_3(\theta) \quad (3)$$

In **Eq.3**, $R_1(\varphi)$, $R_2(\psi)$, $R_3(\theta)$ respectively represents the rotation about X , Y , Z .

- Estimation of posture angle

In order to calculate the rotation angle (φ, ψ, θ) , we need know normal and staring direction of a face. We define the vertical line between nose tip and face as normal direction vector $\overrightarrow{P_{nb}P_{np}}$,the horizontal watch direction as staring direction vector $\overrightarrow{P_{nb}P_{ns}}$ .Take the method presented in [7] to estimate coordinates of points $P_{ns}(x_3, y_3, z_3)$, $P_{nb}(x_1, y_1, z_1)$, $P_{np}(x_2, y_2, z_2)$,as shown in **Figure 7**.

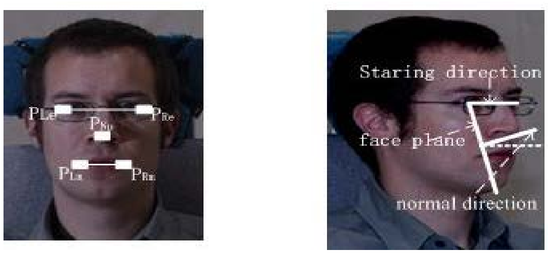


Figure 7. Normal direction and staring direction of a face

Rotation angles are computed as follows.

a) Estimation of φ

Suppose new coordinates of points $P_{nb}'(x_1', y_1', z_1')$ and $P_{ns}'(x_3', y_3', z_3')$ are got by rotating front face about x axis. The correlation can be computed as

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (4)$$

In **Eq.4**, $x_1' = x_1$, $x_3' = x_3$ .After rotating about x axis, the staring direction $\overrightarrow{P_{nb}P_{ns}}$ is parallel to Oxz plane, $y_1' = y_3'$, so **Eq.5** is derived.

$$y_1 * \cos\varphi - z_1 * \sin\varphi = y_3 * \cos\varphi - z_3 * \sin\varphi \quad (5)$$

The rotation angle about x axis is obtained by solving **Eq.5**, coordinates of z_1', z_3' are calculated on the basis of rotation relation.

b) Estimation of ψ

Suppose new coordinates of points $P_{nb}''(x_1'', y_1'', z_1'')$ and $P_{ns}''(x_3'', y_3'', z_3'')$ are got by rotating front face about y axis. The correlation can be computed as

$$\begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix} = \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \quad (6)$$

According to rotation relationship, $y_1'' = y_1'$, $y_3'' = y_3'$, after rotating about y axis, the staring direction $\overrightarrow{P_{nb}P_{ns}}$ is parallel to z axis, $x_1'' = x_3''$ so **Eq.7** is derived.

$$x_1' * \cos\psi - z_1' * \sin\psi = x_3' * \cos\psi - z_3' * \sin\psi \quad (7)$$

The rotation angle about y axis is obtained by solving **Eq.7**, coordinates of z_1'', z_3'' are calculated at the same time.

c) Estimation of θ

After rotating about x, y axes, $x_1'' = x_3''$, $y_1'' = y_3''$, the staring direction is parallel to z axis, finally, the rotation angle about z axis is obtained as **Eq.8**

$$x_1'' * \cos\theta - y_1'' * \sin\theta = 0 \quad (8)$$

2.3. Experiment Simulation

Test our acclimatization calculation framework on ORL human face database. Human face pictures in ORL database are taken by Olivetti Library in University of Cambridge. The ORL human face database contains 40 volunteers' face images, in which each person is regarded as one class. Every class in ORL database has 10 gray images. These face images vary in the facial expression, lighting and posture. We divide 10 images of one person into two parts, one is training set, and the other is testing set. Testing set images are of the following two sorts, every sort has 3 test images and the remaining 4 images are set to be training images.

1) Choose three images with smallest rotation angle (φ, ψ, θ) of the 10, as shown in **Figure 8.**

2) Choose three images with random rotation angle of the remaining 7 images, as in **Figure 9.**

Figure 8. Selected three images with minimum rotation angles

Figure 9. Three images with random angle

After training images and test images are assigned, Principal Component Analysis (PCA) and minimum distance decision are applied for face recognition at last. Recognition rate under these two situations are shown in **Table 1.**

Table 1. Comparison of recogniton rate

Testing set	Image number of training set			
	10	20	30	40
Three images with random rotation angles	90.0%	83.33%	85.56%	75.00%
Three images with minimum rotation angles	96.33%	91.67%	93.22%	85.00%

We can see from the results in **Table 1** that when we take images with minimum rotation angles as test images, it has a great increse in recoginiton rate compared with images with random rotation angles. Results in **Table 1** indicates that face images which have been selected by acclimatization calculation have more advantages on recognition rate than original images, it proves that the proposed acclimatization calculation framework is correct and practical to reduce the effect which is caused posture change.

3. DATA FUSION BASED ON ACCLIMATIZATION CALCULATION

As mentioned above, Multiple Stereovision system can obtain more information about targets than monocular vision. On the one hand ,we can select from multiple images to overcome the difficulty of posture change, on the other hand, images from multiple sensors can be used for data fusion. In Multiple Stereovision system ,all isomorphism or heterogeneous sensor`s redundant and complementary information can be used for data fusion in certain level so as to make holistic assessment about targets more accurately.

3.1. Multi-Sensor Information Fusion

Multi-sensor data fusion technology can be divided into pixel level, feature level and decision level fusion [8].Decision level fusion is to make final decision on the basis of certain principle and reliability of local decision, so there are two basic issues for decision fusion, one is decision principle, the other is how to characterize local reliability. Bayesian Probability Theory, Markovian Process,Dempster-Shafer Evidence Theory are all methods to estimate local reliability. Among them, Dempster-Shafer Evidence Theory is applied more widely due to its ability to describe uncertainty and combination rule. BP neural network is taken to combine with DS Evidence Theory in [9] for information fusion, while RBF is adopted in [10].In this paper, we combine GRNN(General Regression Neural Network) with DS(Dempster-Shafer) evidence theory as fusion method in decision level, shown in **Figure 10**,where GRNN is applied to assign basic probability and DS evidence theory is for decision, every basic probability assigned by GRNN is an evidence for holistic assessment.

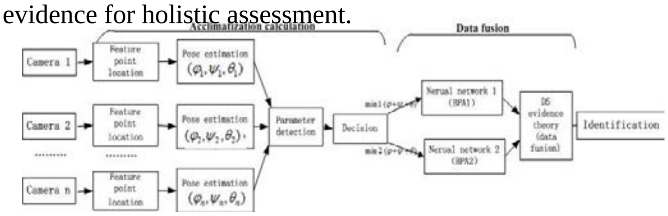


Figure 10. Fusion process based on f acclimatization calculation

3.2. General Regression Neural Network

GRNN(General Regression Neural Network) was proposed by Specht, it is a feedforward neural network which consists of input layer, hidden layer and output layer, see **Figure 11**.

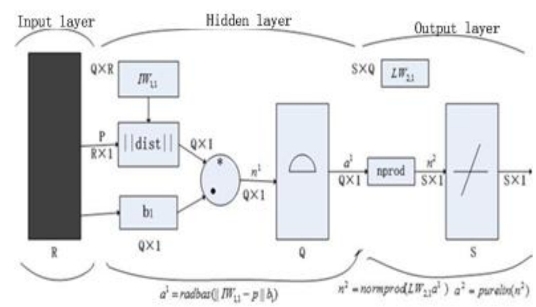


Figure 8. Structure of GRNN

Figure 11.Structure of GRNN

Compare with BP neural network, GRNN has more advantages

1) Train for GRNN is one-way while BP neural network needs iteration

2) Neurons numbers of hidden layer is determined by training samples adaptively

3) Connection weight between layers uniquely determined by training samples while BP neural network modifies the weight in the process of iteration

4) Gaussian function with local activation is adopted as activation function of hidden layer nodes, which has strong attraction to input close to local neurons features.

3.3. Dempster-Shafer Evidence Theory

Dempster-Shafer evidence theory was first put forward by Dempster in 1976 [10], it has the ability to describe uncertainty. In the application of DS evidence theory, the most important step is to estimate BPA (Basic Probability Assignment), definition of BPA is given as follows.

A is an evidence for some proposition, Let Θ be the framework of the proposition's identification, if function $m: 2^{\Theta} \in [0,1]$ satisfies

1) $m(\Theta) = 0$

2) $\sum_{A\subset\Omega} m(A)=1$

Then $m(A)$ is the BPA (Basic Probability Assignment) of A, who stands for the support degree for evidence A.

- Basic Probability Assignment for Face Recognition

Suppose there are n objects faces $F_1, F_2, ...F_n$ in our face library, so their identification framework $\Theta=\{F_1, F_2, ...F_n\}$, store their twelfth Zernike moments [11,12] in library, when selected images after acclimatization calculation of identifying target comes, extract their twelfth Zernike moments correspondingly. Suppose pattern x, y contain m eigenvalues

$x=[x_1, x_2, ..., x_m]$

$y=[y_1, y_2, ..., y_m]$

Euclidean distance between x and y is

$$d=\|x-y\|$$
$$=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+...(x_m-y_m)^2} \tag{9}$$

So the basic probability assignment of evidence i to target F_j is defined as

$$m_i(F_j)=(1-\alpha)\exp(-d_{ij}) \tag{10}$$

α is noise figure, α =0 when there is no noise in system, when identifying target is totally blanked by noise, α =1.Correspondingly, the BPA to uncertainty is

$$m_i(\Theta)=\prod_{j=1}^{N}(1-m_i(F_j)) \tag{11}$$

Before applying DS evidence theory, foregoing assigned BPA should be normalized as **Eq.12**. C is recognition rate of GRNN neural network.

$$m_i = C\bullet y_i \Big/ \sum_{i=1}^{N} y_i \qquad \text{i=1,2...N}$$
$$m(\Theta)=1-C \tag{12}$$

- Combination rule of DS Evidence Theory

Suppose A and B are independent evidences in identification framework Θ, m_1 and m_2 are BPA to A, B, their orthogonal sum is described as

$$m(C)=m_1\oplus m_2$$

New BPA after combination is computed as **Eq.13**

$$m(C)=\frac{\sum_{A\cap B=C} m_1(A)*m_2(B)}{K} \tag{13}$$

$$K=1-\sum_{A\cap B=\Phi} m_1(A)*m_2(B)$$
$$=\sum_{A\cap B\neq\Phi} m_1(A)*m_2(B) \tag{14}$$

K calculated by **Eq.14** is a factor who reflects the confliction degree between A and B. When K exceeds 1, DS evidence theory can not be applied.

- Decision Rules

Before data fusion, certain decision rules should be given, in this paper, we give the following rules.

1) Target category should have the maximum BPA.

2) Difference between target category and other category should be greater than some threshold λ_2.

3) BPA to uncertainty must be less than some threshold λ_1 that is the uncertainty for target category should not be too large.

3.4. Experiment Simulation

There are two parts in our simulation, one is selection part and the other is fusion part. The selection part has been proved in the second section of this paper. And in this section, we only take fusion part into account.

Two images of each person are selected by acclimatization calculation in selection part for information fusion, whose decision rule is $\min 1\,(\varphi+\psi+\theta)$, $\min 2\,(\varphi+\psi+\theta)$ .Then calculate the recognition result according to the information fusion method presented in this paper.

Table 2, **Table 3** and **Table 4** illustrate the fusion process to 3 classes of face in one cycle. M_1, M_2 respectively stand for the selected two images of one person, GRNN 1 and GRNN 2 are general regression neural networks established for the selected two images, and F_1, F_2, F_3 stand for 3 classes of people face.

Table 2. BPA assigned by GRNN 1

Image Name	Class of face images			
	F_1	F_2	F_3	Θ
M_1	0.1763	0.0869	0.0535	0.6833

Table 3. BPA assigned by GRNN 2

Image Name	Class of face images			
	F_1	F_2	F_3	Θ
M_2	0.3564	0.2018	0.1255	0.3164

Conflict factor between M_1 and M_2 is computed based on **Eq.14**

$$K=M_1(F_1)*[M_2(F_2)+M_2(F_3)]+M_1(F_2)*$$
$$[M_2(F_1)+M_2(F_3)]+M_1(F_3)*[M_2(F_1)+M_2(F_2)]$$
$$=0.1294$$

After data fusion, BPA of M_1, M_2 to 3 classes of face is computed as **Eq.13**

$$M_{12}(F_1)=\frac{M_1(F_1)*[M_2(F_2)+M_2(F_3)+M_2(F_\Theta)]+M_1(F_\Theta)*M_2(F_1)}{1-K}$$

$$M_{12}(F_2)=\frac{M_1(F_2)*[M_2(F_1)+M_2(F_3)+M_2(F_\Theta)]+M_1(F_\Theta)*M_2(F_2)}{1-K}$$

$$M_{12}(F_3)=\frac{M_1(F_3)*[M_2(F_2)+M_2(F_1)+M_2(F_\Theta)]+M_1(F_\Theta)*M_2(F_3)}{1-K}$$

$$M_{12}(F_\Theta)=\frac{M_2(F_\Theta)*M_1(F_\Theta)}{1-K}$$

Table 4. BPA assigned after information fusion

Image Name	Class of face images			
	F_1	F_2	F_3	Θ
M_{12}	0.4101	0.6481	0.8003	0.2162

From the data shown in **Table 2, Table 3** and T**able 4**, we can conclude that the fusion method combing DS evidence theory and GRNN not only reduce the uncertainty in face recognition but also improve the accuracy of holistic judgments about target.

Figure 12 and **Figure 13** show the recognition result before and after data fusion . **Figure 12** shows the recognition result of single image with minimum rotation angle in two selected images and **Figure 13** shows the fusion recognition result of two selected images. In these two figures, the X-Coordinate stands for the real class of each test images while the Y-Coordinate stands for the predictive class of each image calculated by neural network.

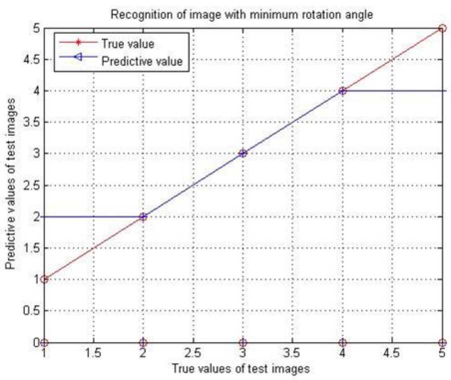


Figure 12. Recognition of single image

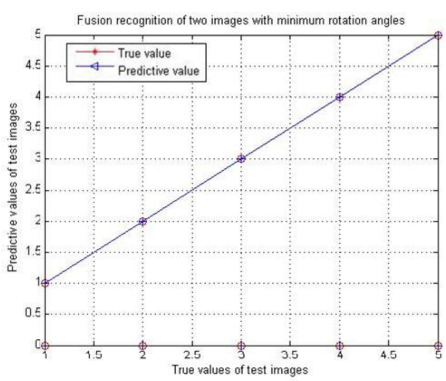


Figure 13. Fusion recognition of two images

From the results, we can observe that there are misrecognitions in **Figure 12**, the face image of class 1 is recognized as class 2, and face image of class 5 is recognized as class 4.However, when we take the fusion result of two images for recognition, all classes are recognized rightly, as is shown in **Figure 13**.

4. SYSTEM DESIGN

Finally, we design a image fusion recognition system on the Matlab platform, as is shown in **Figure 14**.The whole system consists of three main parts, read images, choose recognition method and show the recognition result. For reading images, we will load the images selected by acclimatization calculation which are stored in the local file. Recognition methods are single image recognition and two images fusion recognition. We can choose one recognition method by pressing one of the radio buttons on the interface. At last, we display the matched recognition result.

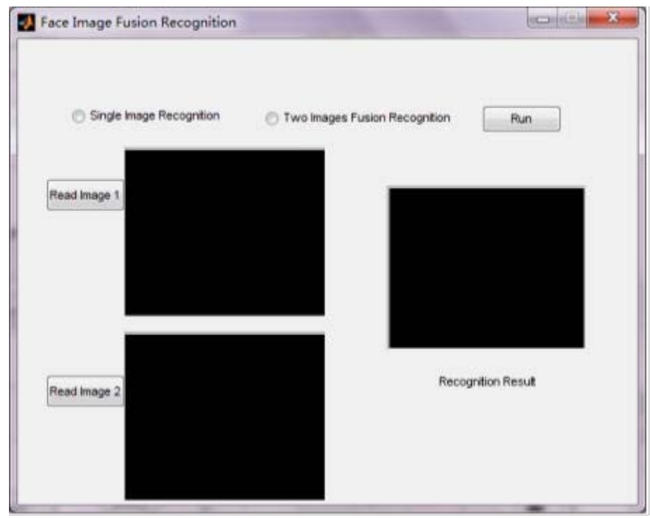


Figure 14. Interface of System

.

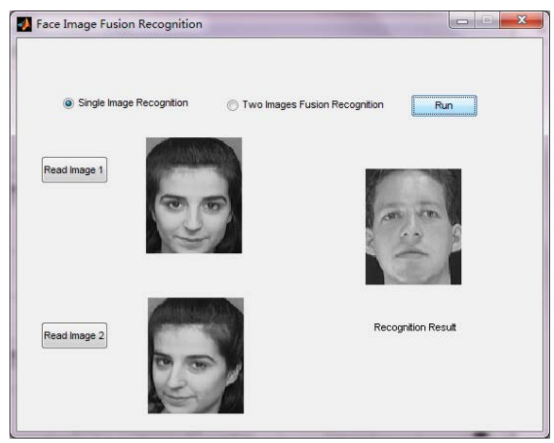


Figure 15. Recognition Result of Single Image

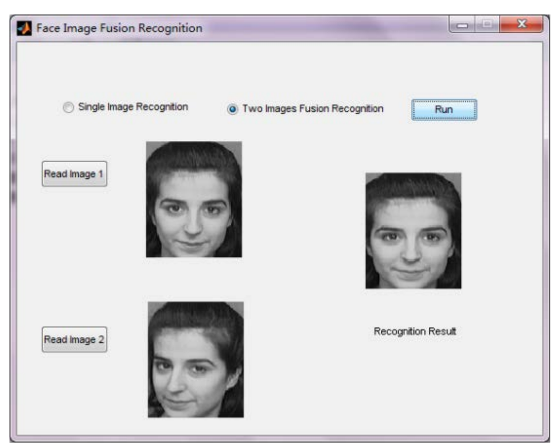


Figure 16.Result of Two Images Fusion Recognition

Figure 15 shows the result of single image recognition about one person while **Figure 16** shows the recognition result of two images' fusion. It can be concluded that when taking two images and applying fusion method in face recognition, it can

reduce the uncertainty and improve the accuracy of recognition. effectively

5. CONCLUSION

The experiment results in this paper have proved that applying Multiple Stereovision in face recognition system and taking acclimatization calculation method to select from input images can reduce the impact of posture change on recognition, in the meanwhile, redundant and complementary information from multi-sensors can be used for data fusion to improve the accuracy of identification. The simulation results prove our thinking to be right and practical.

6. ACKNOWLEDGEMENTS

This work is supported by National Natural Science Foundation(60973060) and Research Fund for Doctoral Program(200800040008)

REFERENCES

[1] Turk M and Pentland A. Eigenfaces for recognition. J.Cogn.Neurosci.3, 1991:72-86.

[2] DaLong Jiang, Yuxiao Hu, Shuicheng Yan, Lei Zhang.Efficient 3D reconstruction for face recognition.J.Pattern Recognition 38(2005):787-798.

[3] Kwang Nam Choi, Marco Carcassoni, Edwin R Hancock. Recovering facial pose with the EM algorithm. J.Pattern Recognition.35 (2002):2073-2093.

[4] Lei Wang, Chao Hu, Jie Wu, Qing He, Wei Liu.Multi-camera face gesture recognition.J.Computer Applications.2011, 30(12).

[5] Guoyuan Liang, Hongbin Zha, Hong Liu. Face pose estimation based on 3D models and affine correspondences.J.Chinese Journals of Computers.2005,28(5):792-800

[6] Cootes T F. Active shape models-their training and application J.Computer Vision and Image Understanding, 1995, 61(1):38-59

[7] Lifang Wu, Sicong Zhang, Xiaoqing Zhao, Ning Wei.A new method for face pose estimation. J.Signal Processing, 2006, 22(1):61-64.

[8] You He, Guohong Wang, Dajin Lu, Yingning Peng.Multi-sensor information fusion and its application. Publishing house of electronics industry, Beijing(2000)

[9] Jin Huang, Yongmei Cheng, Yanni Peng, Quan Pan. Airplane image recognition based on BP neural network and DS evidence theory.J.Computer Simulation,2005,22(11)

[10] Dempster A P.Upper and lower probabilities induced by a multi-valued mapping.J.Ann Mathematical statistics, 1967, 38: 325-339.

[11] S.O.Belkasim, M.Shridhar, and M.Allmadi.Shape-Contour recognition using moment invariants. Proceedings of 10th ICpR.Atlantic City,NJ,pp.649-651,1990

[12] Xiaohong Wang.Reasearch about moments and its application in image processing and recognition. Northwestern Polytechnical University,2001